

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2003, 15:17:50 ; Search time 43 Seconds  
(without alignments)  
1693.016 Million cell updates/sec

Title: US-09-494-297-2  
Perfect score: 3945  
Sequence: 1 MKKTRFPNKLNTLTQRYLS.....IAGISLGIWCIHTIRIKHD 757

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	633.5	16.1	659	2	SA00043
2	603	15.3	638	2	adhesin - Streptoc
3	164	4.2	1039	2	fibronectin-bindin
4	161	4.1	1386	2	protein F2 - Stre
5	159	4.0	1243	2	surface protein (L
6	157.5	4.0	1612	2	sex factor aggrega
7	156	4.0	5005	2	probable peptidogl
8	155	4.0	1530	2	hypothetical prote
9	152.5	3.9	832	2	peptidoglycan anch
10	151	3.9	2013	2	internalin protein
11	149	3.8	1315	2	probable peptidogl
12	149	3.8	1315	2	toxigen B - Clostrid
13	148.5	3.8	1946	2	hypothetical binding
14	148.5	3.8	1883	2	hypothetical prote
15	147.5	3.7	2167	2	hypothetical prote
16	145.5	3.7	1036	2	cell wall-associat
17	145	3.7	1036	2	hypothetical prote
18	144.5	3.7	940	2	S-layer protein -
19	144.5	3.7	1092	2	internalin protein
20	144.5	3.7	1578	2	fibronectin-binding
21	143	3.6	1185	2	peptidoglycan bou
22	143	3.6	1385	2	collagen adhesin -
23	142.5	3.6	627	2	hypothetical prote
24	141.5	3.6	1622	2	prolipoprotein p65
25	141	3.6	3216	2	probable cell surf
26	140.5	3.6	876	2	hypothetical prote
27	139.5	3.5	888	2	hypothetical prote
28	139.5	3.5	1167	2	hypothetical prote
29	139	3.5	621	2	chaperone and heat

30	139	3.5	1433	1	A36734	bacllopeptidase F
31	139	3.5	2366	2	S10317	toxin B - Clostrid
32	138.5	3.5	625	2	T03837	surface lipoprotei
33	138.5	3.5	1582	2	AC1153	adhesin homolog Im
34	138.5	3.5	1983	2	G86643	hypothetical prote
35	138.5	3.5	2013	2	A11489	probable peptidogl
36	138.5	3.5	2484	2	T26215	hypothetical prote
37	138.5	3.5	2607	2	T26215	hypothetical prote
38	137	3.5	1177	2	I64233	hypothetical prote
39	135.5	3.4	3283	2	AC1018	large repetitive p
40	135	3.4	926	2	D86897	hypothetical prote
41	135	3.4	1117	2	S33851	fibronectin-bindin
42	135	3.4	1176	2	A33856	surface-layer 125k
43	135	3.4	1228	2	I40468	surface layer prot
44	135	3.4	1272	2	C90593	hypothetical prote
45	135	3.4	1301	2	S51323	SAC3 protein - yea

## ALIGNMENTS

## RESULT 1

S40043  
adhesin - Streptococcus pyogenes  
C:Species: Streptococcus pyogenes  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 28-May-1999  
C:Accession: S40043  
R:Seila, S.; Avily, A.; Tovli, A.; Burslein, I.; Caparon, G.; Hanski, E.  
Mol. Microbiol. 10, 1049-1055, 1993  
A:Title: Protein F: an adhesin of Streptococcus pyogenes binds fibronectin via two di  
A:Reference number: S40043; MUID:95020565; PMID:7934855  
A:Accession: S40043  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-659 <Seq>  
A:Cross-references: GB:LI0919; NID:9425479; PIDN:AAA26964.1; PID:9425480

Query Match 16.1% Score 633.5; DB 2; Length 659;  
Best Local Similarity 26.1%; Pred. No. 9e-30; Indels 211; Gaps 30;

Matches 203; Conservative 128; Mismatches 237; Indels 211; Gaps 30;

QY	20	SKNSRRFVTLGVFMIFALVTSMVGAKTVGLVESSSTPNAINDSSSEYMYGESYV	79
DB	24	TKRRRFAVTLGVFMIFALVTSMVGAKTVGLVESSSTPNAINDSSSEYMYGESYV	81
QY	80	RGPYKQFRAVHADLVNLEGSRSYQVYCFNLKAPPLGSDSSVKKWKYKKHDKISTKFD	139
DB	82	KEYPGYNIMTRYHDLRVNLNGSRYSQVYCFNIOSNVPSQKNFIRKMFKEIGNGKSFVD	141
QY	140	YAMSPRITGDELNOKLRAVMYNGHPQNMNGIMEGLEPLNATRVQGEAVWYVSDNAPISNP	199
DB	142	YAHYTKLKELEQRLSLLYNAYPNDANGYKKGLEHNAITVTOYAAWHSDNSQYGF	200
QY	200	DESFKESESNLVSTQSLIMKRALQKIDPLAKMPKQVDDQLSTFSESDGDKYN	259
DB	201	ETLWSESEKKEGKISQYTLMEALAKLIDPLELTAANKIPSGRLNIFEE	254
QY	260	KGYONLSGLVPTKPPPGDPPMPNPQPTTSVLRKAYAGYSKLLLEGATL	315
DB	265	EAYONLSLAEVYPPDPKPGE--TSEHNPKTPE	298
QY	316	GDNVNSPFAVRESSDIERIELSDGYTLTVELNSPAGYIAEPITTFVEAGKYTTIDG	375
DB	299	DNLEPTLPVYMLD--GEVY-----PEVPSSELEPALPLMP	337
QY	376	KQIENPKVEI-PYSVEAYNDFEEFSVLTQNYAKFYAKNKNSSQVYCFNADLSP	434
DB	338	QEV--PEKPSIDPLEVPRYE-----FNKKDQSP	364
QY	435	PDSEDGKTYMPDFTTGKVKY--THIAGDLKRYTVKPRDTPDFLKHKKVTEKGRK	493
DB	365	LAGE-----SGETEYITEVYGNO-----QNPVDIDK-----KLPNETGFC--S	399



```

OY 245 ---OLSFESDEKDKYKN-----GYONLSG-----GLVTPKPT 277
D 311 LOGAFVYKTKGRADQSGKELIYOSHSLSFRNMGSYTLTPRNSHVSTNELALSPKSGC 370
OY 278 PGDPMPNPOTTSVLRKAYIGDYKSL-----LEGATLQTLGDNVNSFOARFSSND 331
D 371 SGTSEF--TKPSIVANLKRVAQLRFRKVSITDNPVLPDEAFELKSSNGNS-QKLEASNT 427
OY 332 IGE--RIELSDGYTTLTELNSPAGYSTA-----EPTTF-KVEA 366
D 428 QGELHFKDLTSGYDLKETAPKGYOQVTEKLATVTDVTKRPAEOMKWEKPHSFVXVEA 487
OY 367 GKVTYIID-----GKOIENPKKEIPEVSEAY-----NDPEERSVLTQNYA 409
D 488 NKETIYVNHKEITLFSKKKIMEDDRPDORAKIOVOLLONGQKHPNOIQE---YTKNDW 544
OY 410 KFYV-----AKNKGSSQVYVCFNADLSPDSEGGK-----TMPDPTTGE 452
D 545 SYHFKDLPKYDAKNOEKYKSV-----AEVKVP-----DGKVSYLGNIDFNTRETEFEVEQ 595
OY 453 VKTHHNGRDLFTYVPRDPTFLKHKKVLEKQ--YREKGALEYSGLTQLRA- 509
D 596 NNENLEFGNAEIGQSGPKIIDEDTLTSPKSKIMKNDTAENRPOAI-----OVOLYAD 649
OY 510 --ATQALAYFTDSA---EIDKDKLDYHGFDMNDSTLAVAKILVEYADSNPQLTDL 564
D 650 GVAVEGTRFISGSGNMSFEFKNLKYNGTG--NDLIYSKVEVTPFGD----- 698
OY 565 DFLIPNNKYQSLIGTQWHPEDLVDIRMEDKKEVFPV--HNLTKRTV--TGLADRT 620
D 699 -----VYSSANDIIN-----TKREVITQOGPNLEIEETLPLESAGSGTT 738
OY 621 KDFHEIE-----LNKKOELLSTQVTKD--TNLEK-----GCK-----ATINLKH 661
D 739 -----TVEDSRPVDTLSGLSSEGOQSGDMTIEEDSAHIKFSKRDIDGKELAGTMELRD 793
OY 662 GESLTLOG-----LPEGYSLVKEETDESGYK-----KYNDOEVANATVSKT 703
D 794 SSGGTISTWISDQGVKRFYLMPKYTF-VETAAPDAIEATITFTVNEO--GGVTVNGK 850
OY 704 GITSDETLAEENKKEPVPTG---VDOKI 729
D 851 ATKGDHIVVDAVKPKTSGQVDIRIEKL 880

RESULT 4
AC1533
surface protein (LPXTG motif) [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AC1533
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefelt, J.; Kuhn, M.; Kunsit, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Me
ok, C.; Schluter, T.; Stimes, N.; Tjerstr, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1533
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1386 <GLAY>
A:Cross-references: GB:AL592022; PIDN:GAC96035.1; PID:916413254; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0803

```

Query Match 4.1%: Score 161; DB 2; Length 1386;  
 Best Local Similarity 19.7%: Pred. No. 0.3;  
 Matches 163; Conservative 127; Mismatches 301; Indels 238; Gaps 41;

```

OY 10 LNTLNFORVLKSNK--RFTVTLVGVFLMIFALVTSNMGAKTVF-----GLVE----- 55
D 330 LSTFDTKKVTYMTSMFRSESLDLSISFETNSTETMKLMFGCANLSTIKLKKIKF 389
OY 56 ---SSTPAINPDSSSEYR--WGYEYSVRGHPYKQFRVAHDLRVNLEGRSRYQVCFN 110
D 390 YTVATVEKLVNPPRSAPYNGTW-----TNPKTKRTILSDSLMKVDGATWGYHWS 441
OY 111 LKRAPPLGSDSV--KKWYKKHDCISTKFEDYAMSP---RTGG--DELNOKLRAMYN 161
D 442 AVNNI--IAIDSTISGDWKAEDNFISATDQFMLIDFPLVTVGSVDTTTPGEYITTS 500
OY 162 GHPOANGIMEGLEPLAIRVTOEAVWYYSNAPISNDESEFSKRESNLVSTQSLMR 221
D 501 -----VNGLTFTTIVYKEMQASVVAENSTIYF--KESWK--AEDNPFYSATN---KK 545
OY 222 QALQOLIDPNLATMPKROVPDDFOL-----SIFSEDK---GDKY 258
D 546 GKIDISSVTVTGVDVNTPDYELMYTIDGVSTKIYVKEKSSIEAKDSILYIGDTW 605
OY 259 NK-----GYONLSGGLVPT--KPTPGDPMPNPOTTSVLRKAYIGD 302
D 606 NSKNDFISANDKDNPNVDKIKVEGYNTIKPQTNVYLYXGQSEVITTVK---AD 661
OY 303 YSKLLEGATLQTLGDNVNSFOARFSSNDIGERIELSD---GYTTLTELNSPAGYSIAE 358
D 662 QSTLEADSIITYTDKKNNAKDNFTATDKGNPDPDFDIEGVVDIT-----K 710
OY 359 PITFEVAGKYTTIIDGKOIENPKET-----VEPYSEAVNDEEFSVLTQNYANFY 413
D 711 PGTNKI-----TYIYG---NLSKEVTVYKAAQATLEA--KDSALVGDWTMNSKDNFIS 759
OY 414 AKNKGSSQVYVCFNADLSPDSEGGKIMTPEPT--TGEKYTHHNGRDLFTYVPRD 472
D 760 ATDKDG-----TVDPDKIKVEGTVPTTAGINKVYLYLGNOSKEVTVNKK- 805
OY 473 TDPTPFLKHKKVLEKGYREKGAIEYSG--LTETOLRATQALAIYFTDSALDRDKL 529
D 806 ADQYTL-----EVKDSVITYTGDKMKEDNPISAT-----DKT 837
OY 530 KDYHFGEDMN-----DSTLA--VAKILVEYADSNPQLT-----DLDFIPNNKYQSL 577
D 838 GNSVGFMDIKVEGTVDTKAGTNKVTYTYANOSKEVITVYVADQATLE-----AKDSI 890
OY 578 I--GTQWHPEDLVDIRMEDKKEVFPVTHNLTKRTVYGLAGDTRKDFHEIEL----- 629
D 891 IYTGDKMAED--NFISATDQAGKTIDFKNIKVEGTI---DTTKAGNDITVYSYSGVT 943
OY 630 -KNNKOEELLSTQVTKDNTLEFKDKATINLKHGESLTLQGLPEGYSVLVE-----TDS 683
D 944 RSTLSKTIITYVTKKNQVNLKADST-----LYGDKMKIAKDNFVSATDK 988

RESULT 5
S60138
sex factor aggregation protein precursor - Lactococcus lactis
C:Species: Lactococcus lactis
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 15-Oct-1999
C:Accession: S60138; S49417
R:Godon, J.J.; Jury, K.; Shearman, C.A.; Gasson, M.J.
submitted to the EMBL Data Library, December 1993
A:Description: The Lactococcus lactis sex-factor aggregation gene cluA.
A:Reference number: S60138
A:Accession: S60138
A:Molecule type: DNA
A:Residues: 1-1243 <GOD>
A:Cross-references: EMBL:U04466; NID:9458233; PIDN:AAA66196.1; PID:9458234
R:Godon, J.J.; Jury, K.; Shearman, C.A.; Gasson, M.J.
Mol. Microbiol. 12, 655-663, 1994

```



Db 1410 D-----VTENLTLDASGKLEITNLAPGDIQIETKAPAGYELODTPVVDKITE 1461

Qy 632 NKOELLS-----OTVTDKTNLEFKOKKATINLKHGSLSLQGLPEBYSVLVETDEBG 685

Db 1462 DQKELTLOVTKTNLTNLTGSKVIAEIVDTKGNV-LAEKIHV--GI-VGDKYATAKADIKG 1517

Qy 686 YKV---KYNSQEVANAVSKTGTISDETLAFENKKEPV-----721

Db 1518 YKLKOPNTNKGCVKRETEQK-----TFVEYKKNAPILVVPDPRVTPVKTKPVDPACK 1571

Qy 722 -----PTGVQDKINGIALIVLAGISLGIWJHTIRIK 755

Db 1572 PTVKTSLSPTGDESPYG-----IIFTGLFASFMGFLFKRSK 1608

RESULT 7

F82884

hypothetical protein U0495 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C:Accession: F82884

R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H. submitted to GenBank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit

A:Reference number: A82870

A:Accession: F82884

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-5005 <GLA>

A:Cross-references: GB:AE002148; GB:AF222894; NID:G6899495; PIDN:AAE30907.1; GSPDB:GN001

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: U0495

A:Genetic code: SGC3

Query Match 4.0%; Score 157.5; DB 2; Length 5005;

Best Local Similarity 19.8%; Pred. No. 3.6;

Matches 158; Conservative 100; Mismatches 285; Indels 253; Gaps 36;

Qy 5 RFPKLTMTNLTQRLNSKSRFT-----VTL-----VGFMLTFALVT 42

Db 1598 RIKGTQNLITSPATISODKRYTKCMATNLIALGDIETIVSYTIKNQTKKNVSPITFELPS 1657

Qy 43 SMWGAKT---VFGLVESSTPMAINPDSSSEYRWYGESYVGHYKQFVAHDLRVNLE 99

Db 1658 GPICGYTLAPYFKTISLM-----DQVYHNKMTADIKILR 1694

Qy 100 GSRSYQYVYCNLKAFFLGSDSYKWKYK-KHDSISTKFEDYAMSPRITGDELNOQLRAV 158

Db 1695 -IENIGASLFF-----KDIKFIKRNKDGQISFIHKVS-----1728

Qy 159 MYNHPQANGIMEGLEPLNAIRYTOEAVWYISDNAPISNPDESFKRESNLTSTOLS 218

Db 1729 -----ANDANVEWEFKDLNRREYTLERVY-----LKNKDFNQSINVS-----1768

Qy 219 LMRQALQOLIDPNLATKMPKQVPDDFQLSIFESSEDKDKYKGYONLSGLVTPKPTP 278

Db 1769 -----DYIDLMTIENLNKFK-----LLPTKPLGI 1793

Qy 279 GDDPMPNPQPTTSVLIRKVAIGYSKL-EGATLQLTGDNVNSFOARVSSNDIGE---334

Db 1794 IGAPIKEISDNGAKVQL-KFALNDFDVLKENOTFKF---NIQENNNGANLNEISEHEG 1849

Qy 335 RIELSDG-TYTLTFLNLPAGYSTAEPTTFKVEAGKVTIIDGKQIENPKETIVE-PTSVSE 392

Db 1850 KVELIDQKFEFVANLN-----IKVKEKRYV--KIYFENDQVGVYKIN 1894

Qy 393 AYNDPEE-----FSVLTTONYA---KFYAKKKNSSSOVYVCFNDLKSPPSESGGK 442

Db 1895 FKNDYKRPNNVVDASVNTTQTYFTKFKALASNNLTDVDVANKONISINLDS-----R 1950

Qy 443 TMRPDTTGGVKKYTHIAGRLFKYTVKPRDTPDTFLKHIKKYLEKGYREKGAIEYSG 502

Db 1951 VETIGYHFKAKYISNDNRVWYNTI-PAPTN-----IGNKNNIALNFE-L 1995

Qy 503 TETQL---RAATOLAIYFTTSAEIDKRLKDYHFGDMNSTLAVAKILVEYADOSP 559

Db 1996 NQNDLISRLYTFALY---SKDVNDE-----NANMVYIKNNVPOTISTKP 2042

Qy 560 QUTDDFFIPNNKKVQSLIGTQWHEPDIIDIRMEKKKEVPV-----TNHULTRK 610

Db 2043 SSTYVDLKAQNADEKKTILSLILSHNDI---FEDKNNHLIKAKISIDEIDAHQITNS 2098

Qy 611 TVTGLAGDRTPDFHEIELKNNKQELLSQTVKTDKTNLEFKDGKATINLK-----660

Db 2099 T-----THDYDLTEKENNEMMLTKQLINL-KPNTKRYKKVWFTSKSDTIYNGVN 2149

Qy 661 -----HGESITLQGLPEBYSVLKFEIDSEGYKKVNSQEVANATV--SKTGITSD 708

Db 2150 KDNIVYPIINNHTNTIDLTLEKSPILNSVKT-----KTFNKNESIKILMVGFNKSGSLE 2204

Qy 709 ETLA---FENKKEPV 720

Db 2205 NKYAKLVYKDNNNQOI 2220

RESULT 8

AH1396

peptidoglycan anchored protein (LpxTG motif) [imported] - Listeria monocytogenes (str

C:Species: Listeria monocytogenes

C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AH1396

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioec ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, D.; Jones, L.M.; Karst, U. Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Karapakt, G.; Madueno, E.; Maitournam, A.; Ok, C.; Schlueter, T.; Simoes, N.; Tlherrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of Listeria species.

A:Reference number: AH1077; MUID:21537279; PMID:11679669

A:Accession: AH1396

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1530 <GLA>

A:Cross-references: GB:NC\_003210; PIDN:CAD00654.1; PID:G16412064; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmo2576

Query Match 4.0%; Score 156; DB 2; Length 1530;

Best Local Similarity 20.6%; Pred. No. 0.69;

Matches 160; Conservative 78; Mismatches 221; Indels 316; Gaps 38;

Qy 193 NAPISNDESEFKR--ESESNLVSTOLSIMQALQOLIDPNLATKMPKQVDDFQLSIFE 250

Db 794 NMPISPGEDYDKTGVDESN-----LEYFVKRNENINOSYVIR-----YOTATTL 839

Qy 251 SEDK-----GDKYKNGGYONLSGLVPTK---PPRPDMPNPQPTTSVLIRKAI 300

Db 840 TSDTETTAQIONSNTYTFPGDNTTKGETETKAIKVTIGD---GTGGETKIILNKYDK 896

Qy 301 GDYSKLEGATLQLTGDNVNSFOARVSSNDIGERIELSDGTYTLTFLNLPAGYSI---356

Db 897 ADPSITPLEGATFDLYA-NDEKVDQTTDKNGVIEFDLVGYDYTLKEVASAGETLPTAS 955

Qy 357 AEPTTFKE-----AGVYTTIID--GKQIEN 380

Db 956 TENIQVKLEODEKVVQVWNEKMPIKETGEVHLVKTDRATGATTLAGAERSLYDKGAELO 1015

Qy 381 -----PKKEIV-----EPYSVEAYNDE---EPSVLTONYAKRYVAKN---417

Db 1016 GLTTDEKGLTILHMLDLGSYLLKETKAPAEKYLSEKTIWESVESGQVDAIEIOAENEKDL 1075

Qy 418 -----NGSSQV-----YCF-----427

Db 1076 GEAVLTWVDETNNAKLSGAKFNLLNDSGEVIQTNLVSDENGEIRVONLEPDGVAFOGTEA 1135

```

QY 428 --NADLKS-----PPSEDC-----GKT-----443
      ||:
      ||:
Db 1136 PNNYDLATMTPTTIVAGOTSATMTAENNNKTKPRVDGCVLLVQDSATGTELEGAVF 1195
QY 444 --MTPD-----FTT--GEVKYTHIA-----GRDLFKYTVKPRDID 474
      ||:
      ||:
Db 1196 DLTMTADGAIVASMLTLDANGELITVTLAPGKYSFKETKAPGEGYELATDWEFTIADNPQE 1255
QY 475 -----PPT-FLKHKKYIEGGRKGAIEESGTEQLRAQOLATYFT 519
      ||:
      ||:
Db 1256 KTTTAENTKLAIPAGSFKITIKODESNGVRLAG--AEPSLIAENGTELQTLK---T 1309
QY 520 DSAELDKDLKDYHGFDMNDSTLAVAKILVEYADSNPPQLDLEFFIPNNKQYSLIG 579
      ||:
      ||:
Db 1310 DEK-----GELEVNNLAPGNRI--QETKAP-----DGYO-LBS 1340
QY 580 TQWHPEDLDIIRMEDKKEVIPPVTHNLTRKTV--TGLAGDRTKDFHEIEELKNNKQELL 637
      ||:
      ||:
Db 1341 TPOQFE-----IYANDTSQVTVIAENAKLEPDVAETGAVRLIKTDESETGRLSGAVFSLL 1395
QY 638 SQTVKTKDNLEKDKKATINLKHGESLTLQGLPEGYSLVKTDS-EGYKV-----688
      ||:
      ||:
Db 1396 DESGKVIQANLTTDE-----NGE-IFIDGLTPG-NYSLKETKAPDGEIAEOPWNRQ 1445
QY 689 -----KVNSEVANATVSKTGITS-----ETLAFENKKEP 719
      ||:
      ||:
Db 1446 IYKGVDAVIAIKENSPILANGAISPGEOTDKPESIEIPVKRTDITLATPEVTMLP 1500

```

RESULT 9  
AD1096  
Internalin proteins, probable peptidoglycan bound protein (LPXTG motif) homolog lmo0171

C:Species: *Listeria monocytogenes*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AD1096  
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecher, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Ertlan, K.D.; Fshl, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.; Mok, C.; Schlueter, T.; Simoes, N.; Trieretz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Ertlan, K.D.; Fshl, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AD1096  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-832 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAC98386.1; PID:g16409530; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo0171

Query Match 3.9%; Score 155; DB 2; Length 832;  
Best Local Similarity 19.9%; Pred. No. 0.3; Mismatches 244; Gaps 33;  
Matches 141; Conservative 93;

```

QY 155 LRAVWYNGHPQAN--GIMGELPLNARVTOEAWYYSNAPISNDSFKRESSENV- 212
      ||:
      ||:
Db 176 LKSISSANNKKITGNESLVATLPPELHYLEV-----LGNATITELDITNOENLVY 222
QY 213 -STOSLSMRQALKQOLIDP-----LATKMPQVDDPQSLIFESE 252
      ||:
      ||:
Db 223 LSADELELKLTLKLNLSQNGLRASSISIDMGDESIVLMLPELISDISGNLDSO 282
QY 253 DKGDYKNGYONLISGLVPTKPTPGDPMPPNPQOTTSVLIRKTAIG--DYSKLL--- 307
      ||:
      ||:
Db 283 DHLEMLPAVKMLDISNELTRLPKINDEPL-----LTTINVRSKKIDRLLESSKLVDVP 336
QY 308 -----EGATFLOJT-----GDNNNSFOA-----RVSSNDI-----GER 335
      ||:
      ||:
Db 337 KLATLNADQAVTLKSTIAAGNFTIPNNVENLAGQVTPRTIISNNGTSDQSIAMASGEL 396

```

```

QY 336 IELSDETYLLE--LNSPA-----GYSIAEITPEKVEAGKYVTLIDKQIE-----379
      ||:
      ||:
Db 397 SGLSKVSYTFEDEVINSBALAGKTYGTGNQPIEKA-----VPVIAAKSYAPAVNAKDEA 452
QY 380 -----NPKIEIPEYSVEAYNDFEFSSVLTQNTAKFYAAKNKSGQVYCFNA 429
      ||:
      ||:
Db 453 TFLQDIRASASEMAQITSDYS-----EVDPEATPGDYVTLHAANE-----F 494
QY 430 DKSPDSDSGGCTMPPTTGCVKYTHIAGRLDFYTVKPRDIDPFLKHKKYIEKG 489
      ||:
      ||:
Db 495 DLKA-----DVTYVVAHINDIQKQ 514
QY 490 YREKGAIEY--SGLETOLRA-----ATOL--AIYFTDSAEELDKLKHG--534
      ||:
      ||:
Db 515 VANSNDISPEVGETELTSEVLAKSGAVYTDLDEAKIKNEVDSEYDSSKLTGYEATIIA 574
QY 535 ---FGDMNSTLAVAKILVEYADSNPP--QLTDLFFIPNNKYO-----SLIGIQ 581
      ||:
      ||:
Db 575 KSKSGASSPDKLSVKIV-----DTEKPIIQINNPETIIEKSELREGOIIIDVGITARD 629
QY 582 WHPEDL---VDIIRMEDKKEVIPPVTHNLTRKTVTGLAGDRTKDFHEIEELKNNKQELLS 638
      ||:
      ||:
Db 630 NTDQDLNIMHDSKVDTSK---PGSYEVIT-----YTEDSSGNKSEYTV 670
QY 639 QTVKTKTNLEFKDKRAT-----NLKHGESLTLQGLPEGYSLVKTDSGKYKVN 691
      ||:
      ||:
Db 671 ITVKVEARI---GKITIQYMDSENNELAESITIGGEVETELAKET--EGYTLKEN 724
QY 692 SQ-----EVANATVS-----KITGIDSELTAFENKKEPVYFPGVDCKING 731
      ||:
      ||:
Db 725 PANSGVFEEETROTIIYVKDIINPEFVSENNVPELPSNNNSVNG 774

```

RESULT 10

AD1129  
probable peptidoglycan bound protein (LPXTG motif) lmo0435 [imported] - *Listeria mono*  
C:Species: *Listeria monocytogenes*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AD1129  
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecher, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Ertlan, K.D.; Fshl, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.; Mok, C.; Schlueter, T.; Simoes, N.; Trieretz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Ertlan, K.D.; Fshl, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AD1129  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2013 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAC98514.1; PID:g16409812; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo0435

Query Match 3.9%; Score 152.5; DB 2; Length 2013;  
Best Local Similarity 19.9%; Pred. No. 1.7;  
Matches 169; Conservative 117; Mismatches 302; Indels 263; Gaps 39;

```

QY 18 VLKSNKRPFT---VTLGVFLMIFALVSMGAKTVEGLVSSPTNAINPOSSSEYR-- 71
      ||:
      ||:
Db 1245 LFNMSTREYTLNSVYGVGVAEIIITERSKPTTKI--LDKAVTPNATITIESSDEATEV 1302
QY 72 -----WGYESYVAGHYYKQFRANHLARNLBSRSYQYCFMLKKAFF- 116
      ||:
      ||:
Db 1303 TNEWTLEDSDGVVYTGNTNIR-IPNDEGTYIAKTAATDEAGNTASDEKTPDIDYVPL 1361
QY 117 --LGSDDSVK-----KMYK-----KHDGISTREFDYA-----NSP 144
      ||:
      ||:
Db 1362 LTVYNQDASADEVNSTEANGIKRPLNVAADTDHGNITTPYVDYSKYKMDVLGTYPVYVAT 1421
QY 145 RITGDELNOKLR-----AVVYNGHP-----QANANGIMEG--LE 175

```

```

Db      1422 DASGNKATQVNLRIVDITTSPTLITNNPLFYSIENMKRLAEQELTYAAGLIGDNYDLA 1481
      176 PLNAIRVQEAHVWYSDNAPISNPDESFRK---ESESNIYVSQSLSMRQALQO----- 226
      1482 PGQAVQPNKOPVFTSFSTJFGDIASVAKGOYOVVNLADSSG---NQALPQITITNV 1537
      227 -----LIDENLATKMKQVDP 242
      1538 VDTGVPVIAKADNVSYHVNTTKEAFEPQDARLDVTDNDPDTDLITTSNFAKVLNLRPG 1597
      243 DPQLSIFESEDKGDKYKNGYONLSGLVPTKPPGDPMPBPNOPTTSVLIRYAGD 302
      1598 KVEVITATDTRKGNQOTFEITVQVS---KDKPVITADPKI--SYQGIETVEANFLSGV 1651
      303 YSKLLEGATLQLTGNNVNSFQARVSSNDIGRIELSD-GYTTLELMSNPGAY-SIAPFI 360
      1652 HAEV---TDELDDG-----VAKTSDFAEKVDENKVGTYTVA-LNAKDEYGMTAEV 1698
      361 TPKV---EAGKVVYIIDGKOIENPNKEIVEPYSEAYANDPEEFSVLTTQYAKFYAKN 416
      1699 KVSYSIFKKIAPTEFNNAKNTIEAVN---ELPSLESTFKIE-----AKYLSGN 1744
      417 KNGSSQVYVYCFENADIKSPDSE-----DGGK-----TMTPDFTG-EVKYTHIAG 460
      1745 K---LKVTYTPBQITKGNVPGESYKVTTKDSDGNIAETVTLTIKOTGSPISKMT----- 1797
      461 RDLFKYTYKPRDPTDPTLKHKKYIEKGYREKGAIEYSGLTETQLRAATDIAIYV-FT 519
      1798 KSTKLEVOSEKPEPMNITFFGIKADVIDGDVTKNKIKVDS---EVLNKKVGYTPYFYV 1854
      520 DSAELDKDLKLDYHGFQGMNDSTLAVAKILVEAYDASNPDOLDLDFIPNNKRYQSLIG 579
      1855 DALGESSKLSATQVIVTSPBELTIDKLEISYPIGKT---ISDKQILODIDGKYTNISYG 1911
      580 TQWMEDELVDIIRMEDEKREVIPTVHNLV-----LKRVT-----GLAGDRTKDFH 624
      1912 T-----VKVTNLSKIYDMNKAKYKVTATNNSGGVAEKT----- 1948
      625 FEIELKNNKQELISQVATKDTNLEFK-----DCKATINLKHGSLTLOGLPEYSILV 678
      1949 ILTVYKNSDSSFIAPPSKDDKNNKPAKNIPTGTDLNTELIVMGMLLLVG---GMPFLR 2005
      679 KETDESGYKVK 689
      2006 RKT-----KVK 2011

```

## RESULT 11

S70172

toxin B - Clostridium difficile

C:Species: Clostridium difficile

C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 15-Oct-1999

C:Accession: S70172; S44271

R: von Elchei-Streiber, C.; Meyer zu Heringdorf, D.; Habermann, E.; Sartling, S.

Mol. Microbiol. 17, 313-321, 1995

A:Title: Cloning in on the toxic domain through analysis of a variant Clostridium difficile

A:Reference number: S70172; MUID:96079281; PMID:7494480

A:Accession: S70172

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Cross-references: EMBL:223277; NID:g761713; PIDN:CAA80815.1; PID:g761714

A:Experimental source: Isolate 1470

R: Sartling, S.; von Elchei-Streiber, C.

submitted to the EMBL Data Library, July 1993

A:Description: Cloning and sequencing of an Clostridium difficile toxin B mutant.

A:Reference number: S44271

A:Accession: S44271

A:Molecule type: DNA

A:Residues: 1-1323, 'N', 1325-2367 &lt;SAR&gt;

A:Cross-references: EMBL:223277

C:Superfamily: cpl repeat homology

C:Keywords: cytotoxin

Query Match 3.8%; Score 151; DB 2; Length 2367;

Best Local Similarity 19.9%; Pred. No. 2.7; Mismatches 292; Indels 304; Gaps 42;

Matches 171; Conservative 94

```

      5 REPNNLNTLNTQRYL-----SKNSKRFYTLTGVPFLMFPALVTSWGAQTVGLVE--SS 57
      1568 KPMNRKGSINTSDLSMFLESNMNIKSFVNFQSNIKFLIDANFIISGTSISGQFEFID 1627
      58 TPNAINP-----DSSSEYKMT-GYESTYRGHPYTKQFVAHDLKVNLEGSSYOYVCN 110
      1628 ENNNIQPYFIKENTLETNTVLYGNRQNNIVPENY-----DLDDSDISSTVIN 1676
      111 LKAPPLGSDSSVKKMKYKHGDISKRFEDYASPRTQDELNOKLRAVYNGHPQ---N 166
      1677 FSQKLYLGIDSCVKN-----VVISPNLYTDEINITPYETFNNTYPEYIVLD 1722
      167 ANIGMEGLEPLAIRVTOEAVWYSDNAPISNPDESFKRESNLSVTSQLSMRQALQO 226
      1723 ANYINEKIN-VAINDLISRYVW-----SNDGND-----ILMSTSEKKVQVKIR 1767
      227 LI-----DPLNATKMPKQVPPDFOLSIFESEDKD-----KYNKYQNLISG-- 268
      1768 FVNVFEKDKTLANKLS-----FNSDKQDVPVSEIILSTPSTY---YEDGLIGYD 1813
      269 -GLVPTKPPPTDPPMPBPNOPTTSVLIRKVAIGDIYSKLBGATLQLTGDNVNSFQARV 327
      1814 LGLV-----SLYNEKFYINNEGMVSG--LIYINDSLYKRPV- 1850
      328 SNDIGERIELSDGYTTLTELNSPAGYSIAEPI-----TFKVAGKVV 370
      1851 -NNLTGFTVVDKDKYFNPINGAA-SIGETIIDDKNYFNOGVLQGVSTEDGFTY 1908
      371 -----TI-IDGKOIENPNKEIVEP---YSVEAYNDEEFSVLTTQNY-----AKFY 413
      1909 FAPANTLDEBLEGEALIDFTGKLIDENIYFEDNRYGAVEKELDEGMHYFSPETGAKFK 1968
      414 AKNKGSSQVYVYCFNAD--LKSPPDSEOGKMTDFTTGVEKYTHIAR----- 461
      1969 GLNQIGDOK--YFNSDGYMOKGFWSINKNKYFPDSGYMKAYTEIDKHHYFAENGEM 2026
      462 -----DLFKYTYKPRDPTDPTLKHKKYIEKGYREKGAIEYSGLTETQLRAATQ 513
      2027 QIGVFNTEDGFKY-----FAHNHEDL-----GNEGEIEISYGLNLRNK----- 2066
      514 AIYFTDSAE-----LDKDLKDYHGFQDMN-----DSTLAAVAKILV 550
      2067 -IYFDDSFYAVVGMKDLLEDGSKYYFEDETAAYIGLSLINDGQYFENDGIMQGEVTVI 2125
      551 E-----YAODSNPQLTDL-----FTIPNNN----- 572
      2126 NDKVEYFSDSGIIESGYONIDNRYFIDNGIVQIGVFSDGYRFAFANVNDNIYQO 2185
      573 --KYOSL-----IGTQWMEDELVDIIRMEDEKREVIPTVHNLTRKTVTGL 615
      2186 AVEYGLAVRGVDVYFPGETTYITFGW-----IYDMENESDKYYPFET-----KKAACKI 2236
      616 AGDRTKDFHE-----IELKNNKQELISQVATKDTNLEFRDGKATINLK--HGE 663
      2237 NLIDDIKYYFDEKGIIMRTGLISFENNMYF-----NENGEIQGYINIDBKMFYFGE 2288
      664 SLTLQ-----GLPEGSYLYKE 680
      2289 DGMVQIGVFNTPDGFKRYFAHQ 2309

```

## RESULT 12

T28679

Fibrinogen-binding protein homolog - Staphylococcus aureus

C:Species: Staphylococcus aureus

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-May-2000

C:Accession: T28679

R:Josefsson, E.; McCrea, K.; N. Eklund, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.  
 Microbiology 144, 3387-3395, 1998  
 A:Title: Three new members of the serine-aspartate repeat protein multigene family of St  
 A:Reference number: 220510; MUID:9908700; PMID:9884231  
 A:Accession: T28679  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1315 <JOS>  
 A:Cross-references: EMBL:AJ005646; NID:e1318791; PID:e1318792; PIDN:CA06651.1  
 C:Genetics:  
 A:Gene: std

Query Match 3.88; Score 149; DB 2; Length 1315;  
 Best Local Similarity 20.2%; Pred. No. 1.4;  
 Matches 203; Conservative 93; Mismatches 336; Indels 372; Gaps 45;

QY 8 NKTINTLNTORVLSKNS--KRFYTVLGVFLMIFALVTSMGATVFEVSSRP----- 59  
 DB 208 NSNNENNADILIKRSTAPKRLNTR-----MRIAAYOPSSTEAKNVNDLITSNTLTLYVDA 262  
 QY 60 ---NAINP--DSSEYRWYGESYVRGHPYKQFRVAHDLRVNLEGSRSYQVYCFNLKKA 114  
 DB 263 DKNNKTIYPADYLSLSQITVDKVKSGDYF-----TIKSDTVQVYGLNPFEDI 311  
 QY 115 FPLGS---DSYKKTWKTKHDG---ISTKFEDYAMSPRITGDELNOKLRAVMYNGH--- 163  
 DB 312 KNIGDKIPNNGETIAKHDNTANMLITVFTFDY---DRENSVQMGINYSTYMDA 364  
 QY 164 ---PONANGIMELEPLNAIRVQGEAVWYSDNAPISNPESFRESNL----- 211  
 DB 365 DTFPVSKNDVEFVNTTIGNTTKTANIQIPDYVNEKNSIGSFTEVSHVGNKENPGY 424  
 QY 212 ---VSTQSLSLRQALK---QLIDPNLATKMPKQVPPDFSLFSESDGDKYKNGY 262  
 DB 425 KQTIYVNSPENSILTKAKLKQVAYHSSYPNNIGQINKDVPD---IKIYQ-VPKGYTLNKGY 480  
 QY 263 QNLISGLVPTKPTTGGDPPMPNPOTTSVLIRKAIGCYSK----- 305  
 DB 481 D-----VNTKELT-----DVTNQYLQKITTYGDNNSAVIDFGNADSAVYVNVN 522  
 QY 306 -----LLEGATLQLTGD-----NY 319  
 DB 523 TKQYINSESPILVQMAITLSITGNKSVSTGNALGFTNNOSGAGQEVYKIGNTVWEDTNK 562  
 QY 320 NSFO-----ARVSSN--DIERIELSDGTYYLTLEL-----NSPAGYS 355  
 DB 583 NGVQELGEKGVGNVTYVFDNNTNTFKGEAVTKEDDSYLLIPNLPGDVRVEFSNLPKGYE 642  
 QY 356 IAE----- 358  
 DB 643 VPFSKQGNNEELDSGLSVITVNGKDNLSADLGITYPKYNLGDIYWEDTNKNGIDODE 702  
 QY 359 -----PIPFKEVAGKYV---TIIDGK---QIENPNKEI-----VEPYSEVANDF 397  
 DB 703 KGISGVTVILKDNNGVNLKVTITDADGKYKFTLDNGNKYVEFTTEGPIPTVTYSSDI 762  
 QY 398 EEFSS-VLTQNTY-----AKFYAKNKGSSOVYCFNADLKSPPDSED---GKKT 443  
 DB 763 EKDSNGLTTTGVINGADNMNTLDGCFYKTPKYNLGNVWEDTNKDGRO--DSTEKGISGVY 820  
 QY 444 MPDPFTTGEVYKTHINGRD-LFKYTVKPRDPTDTFLKHKVIEGKRGKGAILEY--- 489  
 DB 821 VTLKNGNEVLQTTKTDKQKYOFT-----GLENGTYKVEFEP 859  
 QY 500 SGLTEQLRAAPOLAI-YYFTDSALDKDK-----LKDY-----HGF 536  
 DB 860 SGITPTQVSSGIDEGIDSGTSTTGVIKDKDNDITDSGFYKPYNLGDIYWEDTNKNGVQ 919  
 QY 537 DMNDSTLAVA-----KIIVEYAQDSNMP--QLTDLDFEIPNNKYSLSIGTQWHPED 586  
 DB 920 DKDEKISGVTVILKDNNGVNLKVTITDDEGKQYQFDL-----NNGTYK----- 963  
 QY 587 LVDIIMEDKKEYIP--VTHNLTLRK-----TVTGL--AGDRTKDFHEIELKNNKQEL 636

DB 964 -----VERETPSGYPTSVTSGNPTREKNSGLTTGTGVIKADNMNTLDSGF---YKTPKYSYL 1016  
 QY 637 -----LSQTVTDKTNLEFKDGKATINLKHESLSLQQLPCEYSYLVAKETSEGGKVK 689  
 DB 1017 GDVYWDYNSKNDGQODSTEEKIGIKDVKKVLLNKEGEVIGTTTDENGKICFDNLDSKRYKI 1076  
 QY 690 VNSQEVANATVSKTIGTSDETLAFENNKKEPVVPTGVDOKI-NGY 732  
 DB 1077 F--EKPRAGLTQGTGNTTDEDKDADGGVVDYTIIDHDFTLIDNGY 1118

RESULT 13  
 AEI149.  
 hypothetical protein lin0132 (imported) - Listeria innocua (strain Cl1p1262)  
 C:Species: Listeria innocua  
 C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: AEI149  
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec  
 D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl,  
 Science 294, 849-852, 2001  
 A:Authors: Krefl, J.; Kuhn, M.; Kunst, F.; Kurparket, G.; Madueno, E.; Maitournam, A.;  
 Ok, C.; Schlueter, T.; Simoes, N.; Tlerraz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla  
 A:Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AEI149  
 A:Molecule type: DNA  
 A>Status: preliminary  
 A:Residues: 1-1946 <GLA>  
 A:Cross-references: GB:AL552022; PIDN:CA093565.1; PID:g16412551; GSPDB:GN00178  
 A:Experimental source: strain Cl1p1262  
 C:Genetics:  
 A:Gene: lin0132

Query Match 3.88; Score 149; DB 2; Length 1946;  
 Best Local Similarity 19.88; Pred. No. 2.6;  
 Matches 185; Conservative 129; Mismatches 359; Indels 262; Gaps 45;

QY 24 KRFTVTLGVFLMIFALVTSMGATVFEVSS---TPN---AINPSSEYRWYGY 75  
 DB 3 KNLLVIVSVLAIASVASFELTRNDAYALIEGFTYSKGLVYFSPAKKNDAMLVN 62  
 QY 76 ESYVRGHPYKQFRVAHDL-----RVNLEGSRSYQVYCFNLKKAFLPGSDSY---KMY 127  
 DB 63 EKTISQKEVKPILNARTFLKDGRIQFLCK---SVALKSASDLIEPSTSYTESKGY 119  
 QY 128 KKHDSIGTKFEDYAMSPRITGDELNOKLRAVMYNGHPOANGIMEGLEPINAL 180  
 DB 120 KATD---TKDAIQAQLPKGTIVKLABGRYIILDNAYIKKNGKLNKKL-PKNTIVSIDENK 175  
 QY 181 -----RVQGEAVWYSDNAPISNPDESFRRE-SESNLVS--TSQSLMRQALKQILD 230  
 DB 176 KVLMEKTELELA---GDQAYIEMENNRKQPLSKSEILVSOQDNEDIDIRISIKYEID 232  
 QY 231 ---NL-ATK--MPKQVPPDFSLFSESDGDKYK-----GYQNL 265  
 DB 233 KAENRMILKATKETTTDTTKDETK---KQEKKNNEOTKTNTNQATNDNTNKAEBANGS 288  
 QY 266 LSGGLVPTKPTTGGD---PMPNPOTTSV---LIRK-----YAGDYK 305  
 DB 289 ANGG--DTGTTTKGNANEGGTTKDNQ--NTTVDKANEIILKLINEAGLTFPOVIVDEL 345  
 QY 306 LLEG-----ATLOLT--GDVNVSRQAVF--SSNDIGERIELSDG---TYLTFLENSPAGY 354  
 DB 346 AVDGQASAKLKITLDSSEKQSLFALILYSEKNVYKVALSSKANQOTDFNSLKGESY 405  
 QY 355 SIAEPTTFKVEACKVYTIIDGQIENPNKEIPEYSEVAYNDPEEFSVLTQYAKF--Y 412  
 DB 406 QVAVGQKYSASQKIDETIFFRQTVAKPVLTPKILIEGENVLTAFELATATELYGKIDEL 465  
 QY 413 YAKNKGSSOVYCFNADLKSPPDSEDG-----GKTMTPD 447



Db 466 VLARKENNSNATSKTYKVDANKLTQNGKVDVDFPDSLNSKEYIIEMEKLVVDGKEYTDD 525  
Qy 448 ---FTTGFKV-YTHAGRDLEFKYTKPRDT-----DPDFIKHIKKV--IEKGRKRG 494  
Db 526 NWYTIASLTAKKPSMAGIDLSYSTEKGFTAVPILDTDNSITNIRYVAILEEDYKANG 585  
Qy 495 QAIE--YSGLETOLRAATOLA-----IYFTDS----- 521  
Db 586 QNAEYAVSYVNASQKKAAYKVRVDMSDGNVFAVAISGNNSNDSFTLATPVSNAVYV 645  
Qy 522 -----AELDKDLK-DYHGF-----GDMNSTLAVAKTLVEADSDNPQ 560  
Db 646 GLSKRPVEFSLEKEADQKLTINYEVEDADNTLLYDNLTHPLTKLYKSDAOGMYGDSFVA 705  
Qy 561 LTDLDFIPNNKKQSLI-----GTQWHPEDVDIIMEKK 597  
Db 706 TVDL---RNKNEITNLLEFGLESEYVVVVMATASYLDDGAGIMVOLLGQSGVFOTK 761  
Qy 598 EVIIVTHTNLTKRVVTGLAGDRTKDFHEIELKNNKQELISQYTKT-DKTN-----LEF 650  
Db 762 EIAVYNAFTLESYTTNA-----EVNKLSEAKKLNSANLKYDKNNLTLYKTYPL 814  
Qy 651 KDGRATINLKHGESLTQGLPEGYSYLVKETDSEGYKVKVNSQEVANATVSKT---GIT 706  
Db 815 NDDENKLTMTEDGYTYLFEELAIKNEYLVKVED--GLDSGMNOVPVEGOLVFETKKASPLT 872  
Qy 707 SDETLAPENKPEVPVPGVCDOKINGYLALEYIAGI 741  
Db 873 DKVLLDYQPNK-----LKVGG-----LAGI 892

## RESULT 14

hypothetical protein U0558 [imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: G82875  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.  
Submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit  
A:Reference number: A82870  
A:Accession: G82875  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1883 <GLA>  
A:Cross-references: GB:AE002154; GB:AF222894; NID:g6899557; PIDN:AAF0971.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: U0558  
A:Genetic code: SGC3

Query Match 3.8%; Score 148.5; DB 2; Length 1883;  
Best Local Similarity 19.1%; Pred. No. 2.7; Mismatches 276; Indels 277; Gaps 34;  
Matches 155; Conservative 102;  
Qy 103 SYOYCFNLKRAK--PLGSDSSVKKWKKHGDISRFPDYAMS----- 143  
Db 951 NFDISMFLKTYFDHPMTSMKRMARITDAMKRYDPFNISKLKIDERANASSVRY 1010  
Qy 144 ----PRITGDELNOK-----LRVYNGHPONANGIMGLEPLNARIYQEAWYYS 191  
Db 1011 KKVLSQILIGEDITNKTFDQFYALLTLQYNASKIWDNKSKEVSKLDELII--EIKKYY- 1067  
Qy 192 DNAPISNDESKFRSESNLYSTQSL-----MRQAQLIDPNIAT 234  
Db 1068 -----DKQFDRGVKNVNVVQTLTFNFDAGIEGNGKYYFKVEKEKELYNILRT 1119  
Qy 235 KMPKQVPD-----DQLSIFSESDGDKYKNGY 262  
Db 1120 KTKESVEIIGIRPSYLTQNKITSYEQLFNKSMLNLFEGITKLIVDGDGDKDNKHLKL 1179  
Qy 263 QNLLSGGLVPTKPTPGDPMPMPNQPTTSVILIRKYYAIG-DYSKILEGATQIOLGDNVNS 321

Db 1180 DELTEKEIYPLMILTPAE-----AAQINFDYKSNNGSFPTLLEN-NG 1222  
Qy 322 FQARFSSNDIGERIEISDGYTTLTFLNSPAGYSIAEPIITFEVAGKYTIID---GKQI 378  
Db 1223 SKFYTIKRDIPSLIEF-----MSVDPK-YTIVENALSSSH 1258  
Qy 379 ENPNK-----EIVEPYSV-----EAYND--FEFESVLT-----ONVAKF 411  
Db 1259 ENVRKWDYELKERDVKFNEVYIKQEAAYKNLLEBFKQULJOLLDGFTESNDIRKF 1318  
Qy 412 YIAKKNGSQYV--CFNADIKSPDSEDGKMTPTPDTGEVYTH-IGARDLTKTV 468  
Db 1319 YKSKDFNPSLDKYKQVDFGKIGLYCFRFSFGKRYERPSPYGEPLDTYVFNAGNPQNVKA 1378  
Qy 469 KPRQTD---POTFLKHKKVIEKGRK-----GQAIYS-GLT----- 503  
Db 1379 KPNQKNVTVDSIINKGIDLEIKRHEINTLNFGLDYAFEFYTYTDI PGSSKDIYMWL 1438  
Qy 504 -----ETQLRAATOLAIYPTDSAEIDKDKIDYHGFGDMNSTLAVAKILV-EY 552  
Db 1439 GSFEGISQSSSEIPNADGIVSSDYTLTKKRQKQELNIEFG-----DYFNIAEVLTRDY 1493  
Qy 553 AQ-----DSNPQDLTDL-DEFIPNNKKYQSLIGTQWHPEDLVIIIMEDKKEYIP 601  
Db 1494 VQTVFIPSONELDNLPNVIISGLSDPNTGNEYVFGSDNTKQNN-----ERLIP 1540  
Qy 602 VTHNLTUKRYTGLAGDRTKDPH-----FEIELKNNK-----QELLSQYV 642  
Db 1541 VNNFLSVNTJIALNTLFAITNNYEKIQNYLANQALNVYQKNOOLFDDYLTLPQELLSNNK 1600  
Qy 643 TDKTNLEFKDGRATINLKHGESLTQGLPEGYSYLVKETDSEGYKVKVNSQEVANATVSK 702  
Db 1601 VKESFLSANPLAL--TTHSDVLANL--NSISFLTRLSPK-----SNSYICK 1645  
Qy 703 TGITSD-----ETIAFENKKEPV 720  
Db 1646 TRLTNNGFKDRWLKRIIDWEIYDNRBPI 1675

## RESULT 15

cell wall-associated protein precursor wapa (B. subtilis) homolog lin0454 [imported]  
A:1489  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AF1489  
R:Glaser, P.; Fraigneul, L.; Buchrieser, C.; Amend, A.; Baguerio, F.; Berche, P.; Bloec  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fshni,  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Simoes, N.; Tietzer, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla  
Ok, C.; Schlueter, T.; Simoes, N.; Tietzer, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AF1489  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2167 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAQ95686.1; PID:g16412895; GSPDB:GN00178  
A:Experimental source: strain Cl1p11262  
C:Genetics:  
A:Gene: lin0454

Query Match 3.7%; Score 147.5; DB 2; Length 2167;  
Best Local Similarity 19.0%; Pred. No. 3.8; Mismatches 319; Indels 331; Gaps 46;  
Matches 179; Conservative 111;  
Qy 47 AKTVFGL-----VESSTPNAINPSSSEYRYGYESYRGHPYKOF-RVAAD--- 93  
Db 562 AGSTFGLNLTPRYLFRVQAVFPGGSPNSDLVFAFMLEK--PQAPVAKAVSNLHKRG 619  
Qy 94 -LRVNLSS--RSYQVCFNLK--KAFPLGSDSSVKKWKKHGDI-STKE---DYAM 142  
Db 620 YVELNMEKSPMADGYKVLVFNKAYEEVDGAEI---KWTQNKGIWPTKEIEIAGKYAL 676

QY 143 -----SPRITGDELNOK-----LRAVMYNGHPONANGIMEGLP----- 176  
Db 677 HHDGKAELAKDPSPYTNSGVNKTERTNYWFRVIAQKAGNNMTTSVQS--EPATPSIPE 734  
QY 177 -----LNAIR---VTOEAVWYISDNAPISNPDESFKRESE-SNL 211  
Db 735 ALNKGQMDVWYTSVPYRGGEVNAATNGNPLFHEHDFNLBGRGPSINVRNRTFNSQDDATGI 794  
QY 212 VSTSLSIMROALKOLIDPNIATKMPKQVPPDDPOLSFESSEDGDKY----- 258  
Db 795 FGKGTSTLEEKLYEENGNIY-----WVESDKKIHFTK--KGDYEAPPGIYSEITKN 847  
QY 259 -----NKGYONLSGGLVPTKPEPTPGDPMPMPNOQOTSVLIRKVAIGDYSKLLE- 308  
Db 848 ADGYLKIIEEDKSETRFLVDGRKSEKDTKGN-----LTVEYTDGKLTSLRDA 895  
QY 309 -GATLOLT--GDNV-----NSFOARVFSNDIGE--RIELSDGTYT---- 344  
Db 896 SGRIVTLTYEGELVKELVGEDRRIKYNDKQELISSYARGLKLYRGYTDGILTSYD 955  
QY 345 -----LTENSPAG-----YSIAEPIYFKVEAGKYTTIIDGKQIE 379  
Db 956 PKHTEEPYETTPAYEEKEKLTETIDPYGKKTTLSTYDKAEQOTTLTNEKKKTTYSYNDAG 1015  
QY 380 NPKKEIYE-----PYSEAYNDEEFSVL--TTONYA-----KPYAKNNKSSQVY 425  
Db 1016 NPKKEIYDADGLKLTYYTESNNLVKENVKPGOEETAYADAGDNITKATDAYGTESYTY 1075  
QY 426 CFNADLKSPDSEDGCKTMTPDFTTGCVKTYTHIAGRDLFKYTYVPRD-----TDPDTFLK 480  
Db 1076 NDNNDYTSSYDTEGRKTYTYD-----GADAVSETLATESQVSSVTOYDAYGN 1123  
QY 481 HIK-----KVIENGYREKGOAIEYSGLTETOLRAA-----TOLA--IYFFDSAEI 524  
Db 1124 PIRGSGELSSGGLNLONGSEFKAGAVSNMTLIQSDAKSMTFDPNTOSAPGALGSGSVKL 1183  
QY 525 DKDKLDYHGFQDMN-----DSTLAVA-----KILVEYQDSNPOLTDL 565  
Db 1184 TSEANSTYKGYSSVTOQVDEPEPTYFSAWIKTSGMTNADALLIGRLQDANKADVTYDAG 1243  
QY 566 F-----FIPNNKYQSLI-----GTQWHPEDLVDT 591  
Db 1244 VMQSNRATSIKKNGDWVKROLTFKTSKNTQVLLYLDNEQAPAPHKGKITWY--DNQOFE 1301  
QY 592 RMEDKKEVLPVTHNLTKRTVTGLAGDRTKDFHEIELEKNNKQELLSQTVKTDKTNLEFK 651  
Db 1302 KGSVASSYNPVVNN-----SEFNHNGTLLPTGMWRTGNTAL--T 1337  
QY 652 DGRATINLKH--GESLTL--QGLPEGYSLVKETDSEGYRVKYNOSQEVANATVSKTGITS 707  
Db 1338 QAKVYNDQSHSGDSAYFEERKATSEATHTHYQD-----VPYNOKRAKALITIS--ALSK 1388  
QY 708 DETLAFENNKPEVPYTGVDOKINGYLALIVYAGISLGIWG 747  
Db 1389 SE-----DAKANGSVATM--SNDYSVMG 1409

Search completed: August 19, 2003, 15:22:29  
Job time : 48 secs